

Some Notes on the Chlorogenic Acids.

3. LC and LC–MS

Version 3 January 2017

Mike Clifford

Emeritus Professor of Food Safety

University of Surrey, Guildford GU2 5XH, Surrey, UK

Contents

	Objectives	
	Summary	
	Referees	
1.	Characterisation by Reversed-phase Liquid Chromatography	Page 4
1.1.	Use of Relative Retention Time or Relative Capacity Factor	Page 4
1.2.	The contribution of equatorial and axial substituents in the (–)-quinic acid moiety	Page 8
1.3.	Chlorogenic acids with one or more <i>cis</i> -hydroxycinnamoyl moieties	Page 12
1.4.	Quinic acid diastereo-isomers	Page 12
1.5.	Chlorogenic acids not containing (–)-quinic acid (L-quinic acid) IUPAC	Page 12
2.	Standards, surrogate standards, and calibrants	Page 16
2.1.	Determining relative retention times	Page 16
2.2.	Calibrants for quantitative studies	Page 17
3.	Chlorogenic Acids Identification and Fingerprinting by LC–MS	Page 20
3.1.	Use of ion-trap-MS ⁿ hierarchical keys for regio-isomer identification	Page 20
3.1.1.	<i>Automated fragmentation of most intense ion</i>	Page 20
3.1.2.	<i>Targeted fragmentations</i>	Page 25
3.1.3.	<i>Distinguishing chlorogenic acid subgroups with identical accurate masses</i>	Page 30
3.1.4.	<i>Discriminating between chlorogenic acids subgroups with accurate masses differing by less than 1 amu</i>	Page 33
3.1.5.	<i>Distinguishing Acyl-quinic Acids not containing (–)-Quinic Acid (L-Quinic Acid) IUPAC</i>	Page 36
3.1.6.	<i>Distinguishing quinic acid diastereo-isomers</i>	Page 37
3.1.7.	<i>Standardisation of Ion-trap-MS protocol</i>	Page 37
3.1.8.	<i>Non-ion-trap MS Instruments</i>	Page 39
4.	Brief Notes on Quinic Acid Biosynthesis	Page 40
5.	References	Page 42

Referees

The following chlorogenic acid researchers are thanked for taking the time to read and comment upon the content of this document before its uploading to Researchgate

Professor Gary Williamson	University of Leeds, UK
Dr Asimina Kerimi	University of Leeds, UK
Dr Dimitrina Zheleva-Dimitrova	Medical University of Sofia, Bulgaria
Dr Edwin Madala	University of Johannesburg, South Africa

Objectives

These notes are designed to draw attention to the complexity of the chlorogenic acids and confusion surrounding their naming, characterisation and identification. They have been placed on Researchgate to make them freely available, and have been reviewed by the undermentioned chlorogenic acid researchers in an attempt to ensure that no errors of fact have crept through, and that the information is presented clearly. Please use them in conjunction with Parts 1 and 2 covering nomenclature and NMR characterisation. Please notify me of any errors that are found and these notes will be updated as necessary.

Summary

Based on this examination of the literature it is recommended that the most efficient method of characterising acyl-quinic acids and structurally related compounds is by LC-ion trap-MSⁿ, using fragment-targeted MSⁿ as required to assist identification and seek minor components that might otherwise be overlooked. Using this approach many acyl-quinic acids can be characterised to regio-isomer level in a single run and without isolation, and without the need for accurate mass. However, accurate mass is helpful in certain situations. Provided that the operating conditions are duplicated, the hierarchical keys prepared by the research groups of Clifford and Kuhnert are rugged and reliable.

Many characterisations are possible with QTOF and triple quadrupole MS instruments, but great care is required to ensure an appropriate collision energy is used to ensure that regio-isomers produce distinctive fragmentations permitting their full assignment. However, these will not necessarily duplicate the fragmentation profiles achieved with ion trap-MS and the hierarchical keys may be invalidated.

Whatever MS is used, it is recommended that the LC-MS system is calibrated with a surrogate standard, such as an extract of a green coffee bean, so that the operating parameters can be adjusted to yield the same pattern of fragmentations as reported by the Clifford and Kuhnert research groups. If this can be achieved, identification of even minor chlorogenic acids is greatly simplified. It is important to ensure that sufficient scans, not less than ten and preferably 20, are

obtained to ensure a reliable fragment profile is obtained, and for minor peaks this might require a sample concentration step.

If the presence of *cis* isomers is suspected, UV-irradiation and / or sodium ion-adduct LC–MS are recommended as recently developed by the Madala research group. If alkyl esters are suspected these can be easily prepared using a surrogate standard in the appropriate alcohol with an acid catalyst. This is a facile reaction and care must be exercised to avoid unwanted production of methyl esters when extracting in methanol.

In situations where full characterisation remains impossible, a reproducible LC–MS fingerprint should still be achievable, and this can be valuable to others who subsequently encounter the same compound.

If MS equipment is not available some characterisation is possible by reference to the relative retention times (relative capacity factors) and even when MS is available these data should be taken into consideration when assigning structures, ensuring that any assignment by MS is consistent with, or at least not greatly inconsistent with, the chromatographic behaviour.

For quantitative studies it is recommended that a good quality commercial 5-caffeoylquinic acid IUPAC be used as the calibrant, its molar absorption value reported, and data expressed as 5-CQA equivalents. This is much cheaper than purchasing small amounts of numerous individual chlorogenic acids and preparing multiple calibration curves, and more importantly, avoids the problem of impure standards generating false calibration curves.

1. Characterisation by Reversed-phase Liquid Chromatography

HPLC, especially reversed phase HPLC, has been used extensively to separate and isolate individual chlorogenic acids. More recently it has been routinely coupled with mass spectrometry, but significant information can be obtained even without MS. This additional information complements MS data and can assist identification.

The separation achieved depends on the column packing and the solvent program, but it is clear that on reversed phase column packings the sequence of elution of acyl quinic acids can generally be explained by the number of free equatorial and free axial hydroxyls in the quinic acid moiety, at least for the extensively studied derivatives of (–)-quinic acid (L-quinic acid) IUPAC.

Often studies report the analysis of a wide range of analytes in a single chromatogram, chlorogenic acids, several types of flavonoids, non-phenols such as terpenes, etc. While there is a need for such analyses and there is a time-saving advantage in an 'all in one' approach, it is not necessarily the best way to assign the structures of any one of the groups of analytes under investigation. It is recommended that the chlorogenic acids, which almost always occur as a complex mixture, be a specific focus so that reliable identification can be achieved as quickly as possible. This will usually be achieved with a simple linear shallow gradient over a relatively long time. Once the chlorogenic acids have been identified, or at least fingerprinted, the gradient system can be revised to speed up the analysis and accommodate other classes of analyte if desired.

1.1. Use of Relative Retention Time or Relative Capacity Factor

The retention time for a compound relative to another is simply the ratio of their retention times expressed as a decimal or a percentage. A better approach is to use the capacity factors rather than the retention times. The capacity factor uses the difference between the retention time and the time to the solvent front leaving the column (time corresponding to the void volume of the column) and is more accurate, especially when the time to the solvent front is large relative to the retention time. However, when examining published data and chromatograms this information might not be available. Both will give the same general picture.

Mono-acyl-quinic acids

When a simple linear gradient is used throughout an analysis, while it may not necessarily achieve the greatest total resolution of a complex mixture of chlorogenic acids and other phenolic and non-phenolic constituents, it does enable the relative capacity factor or relative retention time, to be

used as a structure diagnostic tool.(1, 2) This was once a standard approach but is now often overlooked. Such relationships are illustrated below for a column packing where the 4-acyl-quinic acid elutes before the 5-acyl-quinic acid, and obviously it will be very different for a system where 4-caffeoylquinic acid elutes after 5-caffeoylquinic acid.

	CQA	<i>p</i> CoQA	FQA
3-acyl	0.68	0.68	0.72
4-acyl	0.94	0.91	0.92
5-acyl	1.00	1.00	1.00

The data in Table 1 show that as a first approximation, a 3-acyl-quinic acid and a 4-acyl-quinic acid have a retention time some 30% shorter and some 8% shorter, respectively, than the corresponding 5-acyl-quinic acid.(2) There are relatively few data for 1-acyl-quinic acids, but 1-CQA had a retention time only 50% that of 5-CQA on a phenyl-hexyl column packing.(3)

	Caffeoyl	<i>p</i> -Coumaroyl	Feruloyl
3-acyl IUPAC	1.00	1.25	1.38
4-acyl IUPAC	1.00	1.17	1.27
5-acyl IUPAC	1.00	1.22	1.30

The data in Table 2 show that independent of the position of acylation, a *p*-coumaroylquinic acid elutes some 20% later and a feruloylquinic acid some 30% later than a caffeoylquinic acid.(2) A more extensive data set is presented in Table 3 for a diphenyl column packing.

Similarly, as the hydrophobicity of the acyl residue increases, so does the retention of the corresponding acyl-quinic acid. Accordingly, 3,4-dimethoxy-cinnamoylquinic acids elute after feruloylquinic acids, and are followed by sinapoylquinic acids, and trimethoxycinnamoylquinic acids. In contrast, galloylquinic acids elute before caffeoylquinic acids (and lack the ~320 nm λ_{max}). It is clear from the foregoing that an additional hydroxyl reduces the retention time, and an additional –OCH₃ increases the retention time, and these effects are seen also in the di-acyl-quinic acids.

Table 3. Relative Retention Times for mono-acyl-quinic acids on a diphenyl column packing														
	RT min	Relative Retention Time				RT min	Relative Retention Time				RT min	Relative Retention Time		
		<i>trans</i>	<i>cis</i>	<i>epimer</i>			<i>trans</i>	<i>cis</i>	<i>epimer</i>			<i>trans</i>	<i>cis</i>	<i>epimer</i>
1-CQA	12.0	0.65			1-pCoQA					1-FQA				
<i>Cis-3-CQA</i>	13.2		0.55		<i>Cis-3-pCoQA</i>	18.5		0.64		<i>Cis-3-FQA</i>	20.8		0.65	
3-CQA	13.7	0.74			3-pCoQA	17.2	0.69			3-FQA	21.3	0.76		
<i>3-C-muco-QA</i>	14.4			0.77						<i>3-F-muco-QA</i>	23.5			0.84
5-CQA	18.6	1.00			5-pCoQA	24.8	1.00			5-FQA	27.9	1.00		
<i>epimer</i>	19.5			1.05										
<i>epimer</i>	20.5			1.10										
<i>Cis-4-CQA</i>	21.3		0.89		<i>Cis-4-pCoQA</i>	25.8		0.90		<i>Cis-4-FQA</i>	30.0		0.94	
4-CQA	22.6	1.22			4-pCoQA	28.8	1.16			4-FQA	31.6	1.13		
<i>Cis-5-CQA</i>	23.8		1.00		<i>Cis-5-pCoQA</i>	28.5		1.00		<i>Cis-5-FQA</i>	32.0		1.00	
<i>epimer</i>	26.2			1.41										
<i>epimer</i>	28.2			1.52										

Data recalculated from Jaiswal *et al.* (4, 5)

Di-acyl-quinic acids

For diacyl-quinic acids without C1 substituents on a phenyl-hexyl column the retention times, relative to the dicaffeoylquinic acids as a group, were caffeoyl-feruloylquinic acids 1.19–1.20, diferuloylquinic acids 1.38–1.41, caffeoyl-dimethoxycinnamoylquinic acids 1.38–1.43, and feruloyl-dimethoxycinnamoylquinic acids 1.60–1.66, an increase of approximately 20% for each additional methyl group. Exactly the same 20% increase in retention per added methyl is seen when comparing *p*-coumaroyl-feruloylquinic acids and *p*-coumaroyl-dimethoxycinnamoylquinic acids with *p*-coumaroyl-caffeoylquinic acids on both C₁₈ and phenyl-hexyl column packings.(6)

On a diphenyl column packing caffeoyl-sinapoylquinic acids eluted 30 to 40% later than the diCQA, demonstrating that adding an –OCH₃, as distinct from methylating an existing hydroxyl, also retards the molecule by some 20%. The 3-hydroxy-3-methyl-glutaroyl-caffeoylquinic acids eluted in the period 29 to 44 minutes compared with 33 to 38 minutes for the diCQA indicating a more complex interaction, in part reflecting the hydrophilic character of the aliphatic substituent, but almost certainly also its ability to hydrogen bond and increase the hydrophobicity in some isomers.

The di-*p*-coumaroylquinic acids elute approximately 40% later than the dicaffeoylquinic acids suggesting that the 'loss' of one hydroxyl group has much the same effect as methylating the hydroxyl and keeping the same number of substituents. Inevitably, this results in the diFQA, di-*p*-CoQA, CDQA and CSiQA coeluting, making discrimination in the UV unrealistic if even two sets are present in the same extract.

A similar comparison using a C₁₈ column packing of the retention of the 3,5-diacyl-quinic acid and the 4,5-diacyl-quinic acid relative to the 3,4-diacyl-quinic acid (1.00), are 1.04 ± 0.02 and 1.12 ± 0.02 for nine groups (diCQA, CFQA, diFQA, CDQA, FDQA, di-*p*-CoQA, *p*CoCQA, *p*CoFQA and *p*CoDQA).(6)

There are comparatively few published data for the 1-substituted diacyl-quinic acids. Note, however, that the retention of 1,3-dicaffeoylquinic acid is only ~60% that of 3,4-dicaffeoylquinic acid on C₁₈ and phenyl-hexyl column packings, but the 1,4-di- and 1,5-dicaffeoylquinic acids elute in the same time span as the 3,4-di-, 3,5-di- and 4,5-dicaffeoylquinic acids. Zhao *et al* reported a 1,3-CFQA (tentatively 1C,3FQA) eluting comparatively early from a reversed phase column packing with a retention ~70% that of a 3,4-CFQA (tentatively 3C,4FQA) essentially identical to the retention of 1,3-diCQA relative to 3,4-diCQA (67%),(7) albeit this isomer apparently eluted slightly later than 3,5-diCQA. It therefore seems reasonable to assume that the 1,3-diacyl members of the other diacyl-quinic acid subgroups will also elute earlier, and the structural feature responsible for this behaviour is discussed below.

Yang *et al.* investigating the metabolism of 1,5-diCQA in rats reported four putative CFQA isomers eluting an average of ~40% later and five putative diFQA isomers eluting an average of ~80% later than the substrate but did not fully assign the structures.(8)

If complex gradients are used, the general order of elution might be maintained, but the simple arithmetic nature of the effects will be lost. It is desirable that anyone wishing to utilise such relationships calibrates their own system, and this is essential if a significantly different column packing is employed. Although not expressed in the same format, a similar collection of data are presented by Lin and Harnly, and these illustrate well the effect of varying the

stationary phase.(9) These data are in no way inconsistent with the picture presented here, but do demonstrate clearly that these relationships are not rigid and reinforce the need for careful calibration.

Although such relationships are useful in situations where no MS data are available to assist identification, they are of greater value when checking assignments made on other criteria, such as MS fragmentation, especially, but not exclusively, fragmentations obtained on non-ion trap instruments.

1.2. The contribution of equatorial and axial substituents in the (–)-quinic acid moiety

This section to some extent duplicates the previous section, but examines the data from a different perspective. (–)-Quinic acid (L-quinic acid) IUPAC, as found in the majority of naturally-occurring chlorogenic acids, has two equatorial and two axial hydroxyl groups in its preferred carboxy-equatorial conformer. For a given series of mono-acyl derivatives, e.g. the *trans*-caffeoylquinic acids, 1-CQA and 3-CQA which have two free equatorial and one free axial hydroxyls, elute well in advance of 4-CQA and 5-CQA, which have two free axial hydroxyls and only one free equatorial — for example, see Clifford *et al.*, 2005.(3)

Mono-acyl-quinic acids

The data for 1-CQA are comparatively limited, but it appears mostly to be the first to elute. In contrast, 4-CQA sometimes precedes 5-CQA,(1, 2, 9) sometimes co-elutes (10) and sometimes elutes last,(9, 11-14) clearly dependent on the nature of the column packing.

Note, however, it has been reported that under certain conditions 3-feruloylquinic acid elutes before 1-feruloylquinic acid and 3-isoferuloylquinic acid elutes before 1-isoferuloylquinic acid, but well in advance of the 5-acyl and 4-acyl analogues.(15) Also Zhang *et al.* have reported retention times for all eight caffeoylquinic acids (i.e. both geometric isomers for each of the four regio-isomers) and they report the order of elution as *cis*-3-CQA, *trans*-3-CQA, *cis*-4-CQA, *trans*-4-CQA, *trans*-5-CQA as *cis*-5-CQA, *cis*-1-CQA and *trans*-1-CQA on an ACQUITY UPLC BEH C18 column (2.1 × 100 mm i.d., 1.7 μm with a more complex solvent system and gradient than those used in the previous studies. (A = acetonitrile/methanol 3:1 v/v and B = water containing 0.1% formic acid v/v and a gradient 0–2 min, 2–12%A; 2–5 min, 12–17%A; 5–10min, 17–21%A).(16)

Di-acyl-quinic acids

A similar pattern is observed with the *trans*-dicafeoylquinic acids: 1,3-diCQA elutes first (two free equatorial hydroxyls) and 4,5-diCQA elutes last (two free axial hydroxyls) with 1,4-diCQA, 1,5-diCQA, 3,4-diCQA and 3,5-diCQA (one of each) eluting between them, but closer to 4,5-diCQA, as a fairly tight group. The sequence of this difficult foursome is noticeably dependent on the operating conditions,(3, 9) and any change in column packing requires recalibration with authentic standards or well-characterised surrogate standards.

Tri-acyl-quinic acids

Similarly, one is probably safe in anticipating that 1,3,4-triCQA and 1,3,5-triCQA (one free equatorial hydroxyl) elute before 1,4,5-triCQA and 3,4,5-triCQA (one free axial hydroxyl) but there are few reliable data on which to predict the order within the two subgroups.(9, 17, 18) Such data as are available have 1,4,5-triCQA eluting before 3,4,5-triCQA,(9) and 1,3,5-triCQA before 1,3,4-tricaffeoylquinic acid.(19)

Tetra-acyl-quinic acids

Tetra-acyl-quinic acids have no free hydroxyl group on the quinic acid moiety. There are two distinct subgroups:

- (i) those containing one or more aliphatic residues with a balance of hydroxycinnamoyl residues, and
- (ii) those containing only hydroxycinnamoyl residues.

1,3,4,5-Tetra-caffeoylquinic acid is known,(20-23) but 1,3,4,5-tetra-galloylquinic acid is the best characterised of the second group, eluting a little later than 3,4,5-tri-galloylquinic acid.(24) The behaviour of those in group one is complicated by the potential for the distal carboxyl of one or more of the aliphatic residues to hydrogen bond and no simple relationships can be defined at present.(25, 26)

Depsides

The best known depsides are galloylquinic acid derivatives where one or more galloyl residues esterify one or more gallic acid residues attached directly to the quinic acid moiety. The retention time increases for each additional depsidic galloyl residue, but pentagalloylquinic acids with a 1,3,4,5-tetragalloyl core elute before pentagalloylquinic acids with a 3,4,5-trigalloyl core, i.e. two depsidic galloyl residues increase the retention to a greater extent than one depsidic galloyl residue plus another attached directly to C1 of the quinic acid.(24)

Depsidic shikimic acid derivatives (di-, tetra-, penta- and hexa-galloylshikimic acids) also are known,(27, 28) and may have been overlooked.

In so far as data are available, it appears that these guidelines presented in the various categories of chlorogenic acids discussed above are applicable to chlorogenic acids containing *trans*-hydroxycinnamic acids other than caffeic acid,(6, 14, 24, 29) but not to their alkyl esters, e.g. methyl caffeoylquinates, or quinides. There are no relevant data for hydroxybenzoic acids other than gallic acid.

Methyl chlorogenates

The methyl caffeoylquinates elute from a diphenyl column packing in the sequence methyl 3-CQ, methyl 1-CQ, methyl 4-CQ and methyl-5-CQ, rather than 1-CQA, 3-CQA, 5-CQA and 4-CQA. This revised order was also observed

for the methyl FQ and methyl *p*CoQ. In contrast the methyl diCQ, methyl diFQ, methyl CFQ, methyl *p*CoCQ and methyl *p*CoFQ elute in the same order as the free acids, but note that the only one member of each pair of the mixed diacyl methyl quinates was available for investigation. (30)

Mixed di-acyl-quinic acids

When the quinic acid moiety bears two different hydroxycinnamoyl substituents, for example one caffeoyl and one feruloyl moiety, the data available suggest that the 3,4-diacyl pair precede the 3,5-diacyl pair, and that the 4,5-diacyl pair elute last. There are few data for the 1-acyl analogues but re-examination (further discussed below) of data from Han *et al.*(31) suggests that 1,5-CFQA elute between the 3,4-CFQA and the 3,5-CFQA. The members of each pair have the same number of free axial and equatorial hydroxyl groups in the quinic acid moiety, and the same number of aromatic hydroxyls and OCH₃ groups in the acyl moieties, and some other feature must explain their resolution.

Although there are relatively few data from which to generalise it appears that if the more hydrophilic acyl residue is attached equatorially it will elute before the other member of the pair,(6, 9) i.e. 3F,4CQA elutes before 3C,4FQA, and 3F,5CQA before 3C,5FQA.

This relationship has also been observed by Clifford *et al.* for the 3,4- and 3,5-*p*CoCQA,(6) and is supported by data from Lin and Harnly, except that in their study the two 3,4-*p*CoCQA could not be resolved.(9)

It is well established that *p*-CoQA elute before the FQA, i.e. that *p*-coumaric acid is more hydrophilic than ferulic acid, and on this basis the relationship holds also for the 3,4- and 3,5-*p*CoFQA.(13)

However, for the 4,5-diacyl pair of mixed isomers both acyl groups are attached to equatorial hydroxyls. Such observations as are available indicate that the first of the pair to elute has the more hydrophilic substituent at C5, e.g. 4F,5CQA elutes before 4C,5FQA, 4*p*Co,5CQA before 5*p*Co,4CQA, and 4F,5*p*CoQA before 4*p*Co,5FQA.(6, 9, 13) As discussed in Part 2 of these notes the diaxial ³J coupling constants reported by Pauli *et al.* suggest some flattening of the cyclohexane chair with the H_{4ax}-H_{5ax} dihedral angle as judged from the Karplus relationship nearer 140° rather than 180°. Tolonen *et al* using ¹H-NMR data and the Altona-Hasnoot equation calculated the H_{4ax}-H_{5ax} and H_{5ax}-H_{6ax} angles in the range 148–163° for five mono- and di-acyl-quinic acids.(32) Possibly such distortion of the ideal chair conformation renders the acyl moiety at C5 more equatorial than that at C4, and thus effectively extends the relationship, first reported by Lin and Harnly, also to the 4,5-diacyl mixed isomers.

However, a very different pattern, with many more fragments, is presented for the CFQA in a study by Han *et al.*(31) They did not cite the studies by Clifford *et al.* or use their hierarchical keys,(3, 14) and used an ionisation voltage of 4.5 kV rather than 3.5 kV, and did not define the collision energy employed. These factors no doubt account for the different behaviour with more fragments and difficulty of assigning the individual isomers. Several of their assignments are doubtful because of significant discrepancies in the reported sequence of elution, for example 3-

CQA after 5-CQA, 3F,4CQA immediately before 4F,5CQA with 3C,4FQA last of this group of eight CFQA. **Operating conditions must be standardised if the hierarchical keys from the Clifford and Kuhnert research groups are to be used.**

The (3-hydroxy-3-methyl-glutaroyl)quinic acids elute very early, in the period 4 to 6 minutes from the diphenyl column packing relative to ≈ 19 min for 5-CQA. It is therefore not surprising that a mixed diacyl-quinic acid containing one or more aliphatic dicarboxylic acid moieties also elutes comparatively early, e.g. the (3-hydroxy-3-methyl-glutaroyl)-caffeoylquinic acids overlap with the elution of the dicaffeoylquinic acids not acylated at C1, the sequence of elution for the six regio-isomeric (3-hydroxy-3-methyl-glutaroyl)-caffeoylquinic acids is very different to that seen for the mixed diacyl-quinic acids discussed above, for example see Clifford *et al.*,⁽³³⁾ probably reflecting the ability or otherwise to form internal hydrogen bonds, and further generalisations cannot be made.

Chlorogenic acid glycosides

Glycosides of chlorogenic acids, particularly caffeoylquinic acid glycosides and dicaffeoylquinic acid glycosides are increasingly reported. The hydrophilic sugar moiety reduces the retention time considerably, for example 5-CQA-4'-glycoside and 5-CQA-3'-glycoside, both elute before 5-CQA,⁽³⁴⁾ with retention times relative to 5-CQA of 0.72 and 0.95, respectively. Dicaffeoylquinic glycosides elute from a phenyl hexyl column packing earlier than 1,3-dicaffeoylquinic acid, the most hydrophilic of the six dicaffeoylquinic acids.⁽¹⁷⁾ Because of the marked impact of glycosylation position on the retention time, the sequence of elution for the eight possible caffeoylquinic acid glycosides, and 24 dicaffeoylquinic acid glycosides, is complex and not fully established.⁽¹⁷⁾

Ma *et al* using a C₁₈ column packing reported very hydrophilic caffeoylquinic acid diglycosides (retention time 0.27 to 0.82 relative to 5-CQA), caffeoylquinic acid glycosides (0.60 to 1.60) and dicaffeoylquinic acid glycosides (1.54 to 2.14).⁽³⁵⁾

Reports of putative dicaffeoylquinic acids eluting before 5-caffeoylquinic acid IUPAC are very likely to be caffeoylquinic acid glycosides, and the following reports are likely to be incorrect assignments.⁽³⁶⁻³⁸⁾ Similarly, a report of a putative caffeoyl-diferuloylquinic acid which eluted in advance of the diferuloylquinic acids is more likely to be a diferuloylquinic acid glycoside.⁽³⁹⁾

Baeza *et al.* have reported in *Coffea arabica* novel caffeoyl-2,7-anhydro-3-deoxy-2-octulopyranosic acids which can be distinguished from the coexisting dimethoxycinnamoylquinic acids by accurate mass ($m/z = 381.0820$ and $m/z = 381.1186$, respectively).⁽⁴⁰⁾ Baeza *et al.* were not able to provide comprehensive fragmentation data but it is unlikely that the novel glycosides would produce a substantial m/z 207 that is characteristic at MS² for the dimethoxycinnamoylquinic acids.⁽²⁹⁾ Data are available for the fragmentation of the *p*-coumaroyl, *p*-hydroxybenzoyl and benzoyl analogues reported in *Erigeron*.⁽⁴¹⁾

immediately allow the possibility that one of the commoner chlorogenic acids has been wrongly identified. Similarly, there has never been any suggestion that more than one or two novel chlorogenic acids are present, again possibly raising questions about the reliability of the assignment, but more importantly highlighting the need for a specific method of analysis to seek what might be minor components eluting in a complex and crowded part of a chromatogram.

Acyl-epi-quinic acids

Be that as it may, it was concluded in Part 2 that there was convincing evidence for the natural occurrence of at least one *epi*-quinic acid derivative based on the significantly different chromatographic behaviour of (–)-quinic acid (L-quinic acid) IUPAC when compared with the product released by saponification of the novel chlorogenic acid. The compound isolated by Wang *et al.* appears to 4,5-dicaffeoyl-*epi*-quinic acid IUPAC. 3,4-DiCQA IUPAC was not detected, but 3,5-diCQA IUPAC and 4,5-diCQA IUPAC were present, eluting later than the novel compound. The early elution of the novel compound is unexpected for an isomer which in the carboxy-equatorial conformation would have only two free axial hydroxyl groups, and this behaviour suggests that it chromatographs as the carboxy-axial conformer, the conformation favoured by *epi*-quinic acid with two free equatorial hydroxyl groups — See Part 1, Table 5.

Taking the novel chlorogenic acid as the anchor, and calculating the relative retention times from visual inspection of an enlarged copy of the published chromatogram, gives relative retention times of 1.00, 1.03 and 1.12, identical to the relative retention times for the 3,4-di-acyl, 3,5-di-acyl and 4,5-di-acyl(–)-quinic acids in nine chlorogenic acid subgroups as discussed above. While this must raise some concern about the reliability of the identification, the distinctive NMR data and demonstrated presence of a quinic acid moiety other than (–)-quinic acid must be given greater weight. It does, however, demonstrate how complex a mixture might be present in this portion of the chromatogram, and reinforce the need for a specific method to seek out such chlorogenic acids.

Incompletely characterised epimers

In contrast to the foregoing there are records of four incompletely characterised caffeoylquinic acids, albeit minor components, that are accompanied by a comprehensive set of the commoner caffeoylquinic acids, and are thus chromatographically distinct from the *trans* and *cis* isomers thereof.

Two of these with distinctive fragmentations have been observed in maté, eluting later than *cis*-5-CQA, and are thus appreciably more hydrophobic.(47) Two more have been observed in certain Asteraceae, eluting between *trans*-5-CQA and *cis*-4-CQA, also clearly distinguished by a distinctive MS² base peak,(5) as presented in Table 13.

Synthetic acyl-muco-quinic acids

Data from Jaiswal *et al.* for synthetic 3-caffeoyl- and 3-feruloyl-*muco*-quinic acid record their elution approximately 5% later than 3-CQA and approximately 10% later than 3-FQA, respectively, and well in advance of 5-CQA and 5-FQA, respectively. These two acyl-*muco*-quinic acids have two free equatorial and one free axial hydroxyls and an equatorially disposed hydroxycinnamoyl moiety in the favoured conformer, and might therefore have been expected to elute even earlier. It follows from the retention time that 3-caffeoyl-*muco*-quinic acid cannot be one of the four incompletely characterised caffeoylquinic epimers referred to above.

Anomalous chromatographic data

The chromatographic behaviour of chlorogenic acids described in the foregoing, with minor variations dependent on the nature of the column packing and solvent gradient, can be found in hundreds of publications.

However, there are several data sets where the reported chromatographic behaviour is completely different. These have been obtained by a South Korean research group using a Shiseido (Chuoku) Capecll Pak C₁₈ column with 0.05% phosphoric acid in water for solvent A and methanol for solvent B, and an isocratic step gradient, as follows: 40% B isocratic from 0–10 min, 50% B isocratic 10–20 min, 60% B isocratic 20–30 min and 40% B 30–35 min.(52-59)

The reported elution order is 3,4-diCQA eluting at 2.7 min (before the mono-acyl quinic acids, 3-*p*CoQA at 2.9 min, 5-CQA at 4.5 min, 3-CQA at 8.9 min, with 3,5-diCeQA at 9.6 min, 3,5-diCQA at 12.9 min and 4,5-diCQA at 17.1 min. Nugroho *et al.* describe their standards as having been supplied by Professor Kang Ro Lee (see Kwon *et al.* (60)) and in this original account there is no reference to 3,4-diCQA. In some publications 3,5-diC-*epi*-quinic acid is referred to as 3,5-diC-*muco*-quinic acid,(55-57) and what appears to be 3-*p*CoQA is referred to as '3-O-*p*-coumaroyl-caffeoylquinic acids',(53) possibly a typographical error, but further confusing the accounts.

The assignment of 3,5-diCeQA, or 3,5-diC-*muco*-QA, as discussed elsewhere, is considered doubtful. That constraint, coupled with the idiosyncratic analytical procedure make it difficult to relate their otherwise interesting data to those obtained from the bulk of the literature.

2. Standards, surrogate standards, and calibrants

In this section the term 'calibrant' is used in two contexts. The first is to provide relative retention times or relative capacity factors for which either expensive commercial standards or relatively cheap surrogate standards,(35) or mixture thereof, can be used. Surrogate standards would be well characterised preparations such as an in-house generated extract of a green coffee bean. The second context is as a standard for quantitative studies.

2.1. Determining relative retention times

Calibration of a chromatographic system so as to enable relative retention times or relative capacity factors to be used, requires substantial data sets covering most, if not all, members of several chlorogenic acid subgroups.

5-Caffeoylquinic acid IUPAC and 1,3-dicaffeoylquinic acid IUPAC are readily available and quite cheap, and can be used as the anchor for such calibrations. Many other chlorogenic acids are listed by chemical suppliers, but often with confusing and sometimes incorrect nomenclature as discussed in Part 1, and usually at a substantial price for milligram quantities that frequently have significant impurities. 1,5-DiCQA IUPAC may be worth purchasing if 1,4-diCQA is required because it is a better starting material than 1,3-diCQA from which to produce that isomer by acyl migration.

For many chlorogenic acids a cheaper and often much more convenient approach is to use surrogate standards, in combination with acyl migration in base,(1, 2, 9) partial acid hydrolysis,(3, 9) and UV irradiation (42) applied to a limited set of commercial standards (or in house isolates) to manufacture a fairly comprehensive set of standards.

Acyl migration of 5-caffeoylquinic acid in either ammonia or tetramethylammonium hydroxide will easily provide the 3-caffeoylquinic and 4-caffeoylquinic acids, and mild acid hydrolysis of 1,3-dicaffeoylquinic acid will additionally provide 1-caffeoylquinic acid.

An extract of a green arabica coffee bean will provide 3-CQA, 4-CQA and 5-CQA, 3-FQA, 4-FQA and 5-FQA, plus 3,4-diCQA, 3,5-diCQA and 4,5-diCQA. Almost always 5-CQA, 5-FQA and 3,5-diCQA are the dominant members of their respective subgroups. There is a risk that 3-FQA and 4-FQA might coelute with other chlorogenic acids and not be locatable unless MS is available. Although 5-*p*CoQA is usually detectable in an extract of a green arabica coffee bean the 3- and 4-*p*CoQA isomers may not be and a better source of these is fermented, i.e. alcoholic, apple juice, variously known as cider in the UK or hard cider in some countries. Note that rather unusually 4-*p*CoQA and 5-CQA dominate their respective subgroups and this has often led to 4-*p*CoQA being described incorrectly as 5-*p*CoQA.

A green robusta coffee bean can be used to locate many other minor chlorogenic acid subgroups not present in the arabica if MS detection is available. If green coffee beans are not available an instant coffee can be used for the CQA, FQA and diCQA, but MS detection is probably essential because of the more complex chromatogram. This is also a good source of caffeoyl- and feruloyl-quinides, and potentially 3-caffeoyl-*muco*-quinic acid and 3-feruloyl-*muco*-quinic acid albeit these are minor components formed during roasting.

2.2. Calibrants for quantitative studies

A calibrant for quantitative studies must be of high purity. A commercial standard described as 'chromatographically pure' may simply mean that at a chosen wavelength, probably 320 nm for chlorogenic acids, only one substantial peak will be seen in the chromatogram. An examination at 280 nm might reveal many other components. In addition there might be ash, water and / or organic solvents present, none of which would be detected by chromatography with UV detection. Any of these impurities will contribute to the mass weighed when preparing a calibration curve and thus result in a more dilute solution than expected.

Early studies on chlorogenic acids where a novel compound was first reported almost always involved relatively large scale preparations yielding grams of the material, sufficient to permit repeated recrystallisations, drying *in vacuo*, proximate analysis including ash content, etc. The final preparation would then be characterised by melting point, specific rotation, UV-vis spectrum and molar absorbance, possibly plus IR and NMR spectra if such procedures existed at that date. The collected molar absorbance values make interesting reading (61, 62) — see below. For a given acyl residue, mono-acyl regio-isomers differ by no more than approximately $\pm 4\%$ of mean, and by no more than $\pm 6\%$ for the dicaffeoylquinic acids.

Accordingly one would expect the slope of calibration curves (response factors) for any set of mono- or di-acyl-quinic acids to differ by no more than $\pm 6\%$ of mean. Data from Ruan *et al.* for 3-CQA, 4-CQA and 5-CQA IUPAC do indeed meet that expectation.(63) However, compare the other four sets of data which show substantial variations up to approximately three-fold,(53, 64-66) and data from Liu *et al.* and Che *et al.* (not shown) provide further examples.(67, 68) Clearly, compositional data obtained by such calibration curves will be unreliable, and might distort which isomer dominates.

It is simpler, and cheaper, to use a good quality 5-CQA IUPAC standard and express all data as 5-CQA equivalents, with a correction if desired for the diacyl-quinic acids. If the molar absorption value for the 5-CQA is provided others can then properly compare their quantitative data.

Accurate quantification requires also the avoidance of artefacts during sample preparation. The potential for acyl migration, partial hydrolysis, alkyl ester formation and addition of water across the cinnamic acid double bond, are well documented — for example see Clifford (1985),(69) Deshpande *et al.* (2014)(70) and Wianowski *et al.* (2015).(71)

3. Chlorogenic Acids Identification and Fingerprinting by LC–MS

For the purposes of this exercise the MS equipment can be divided into two types — ion-trap instruments, and other types. Some of these are so-called accurate mass, others accurate only to unit mass, and while access to an accurate mass instrument is useful, an ion-trap that can only achieve unit mass accuracy is more powerful than an accurate mass non-ion-trap instrument for determining chlorogenic acid regio-chemistry. LC–ion-trap analysis is the gold standard for chlorogenic acids characterisation, permitting all chlorogenic acids in an extract to be fingerprinted and most fully identified in a single analysis without a need for isolation. Minor components can be accommodated easily, either by a larger injection volume or pre-concentration, and usually, overlapping peaks can be resolved in the fragmentation spectrum even if not chromatographically.

3.1. Use of ion-trap-MSⁿ hierarchical keys for regio-isomer identification

Ion-trap instruments are designed to ionise the molecules of interest, select ions, i.e. trap ions, of a particular m/z ratio, and if required cause them to fragment and record the m/z values of these daughter ions that form the MS² spectrum. If desired one of these daughter ions can be selectively trapped and fragmented to produce the MS³ spectrum, and this can be repeated several more times provided the signals are sufficiently strong. For chlorogenic acids negative ion operation is to be recommended — structural information for chlorogenic acids is much more difficult, if not impossible to obtain, in positive ion, and it is generally false economy to try and fully characterise flavonoids and chlorogenic acids in one run.

Having selected negative ionisation, defined fragmentation conditions and other operating parameters, an ion trap can be operated in two distinct modes. The first is an automated system where the instrument attempts to ionise everything in a selected m/z range (e.g. 100–800 m/z), fragments everything to produce numerous MS² spectra, automatically selects the strongest ion in each MS² spectrum, and repeats the fragmentation step to produce MS³ spectra, and so on. This can be a useful general screening procedure, and sometimes is all that is required to give a full identification.

The second mode of operation is much more targeted, provides more detailed information that is often essential for identification of the regio-isomer, but is focussed on a particular chlorogenic acid family, or even a single chlorogenic acid. The following examples illustrate the relative merits of these two approaches.

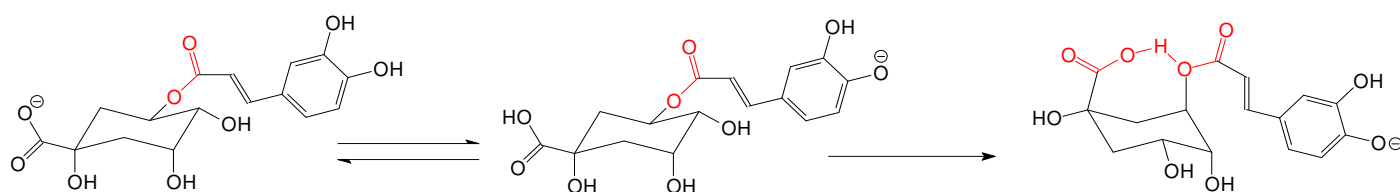
3.1.1. Automated fragmentation of most intense ion

The following table summarises in a simplified format the fragmentation data most important for assigning regio-isomers — full spectra and precise operating conditions, are given in the paper describing the original hierarchical key.⁽¹⁴⁾ The various hydrogen bonding networks that facilitate the distinctive fragmentations are illustrated below.

cis-5-acyl-quinic acids being assigned as *trans*-1-acyl-quinic acids.(42) Another approach to characterise *cis* isomers is to use sodium-ion adducts,(48) as discussed below.

Note that investigations using a quadrupole ion-mobility time-of-flight mass spectrometer have established that *trans*-4-caffeoylquinic acid and both *cis*- and *trans*-5-caffeoylquinic acid produce ions with the charge located on the quinic acid carboxyl as well as on the phenolic hydroxyl. The four prototropic isomers derived from 5-CQA (two *cis* and two *trans*) are resolved in the ion mobilograms. Each pair of prototropic ions differ slightly, but distinctly, in their fragmentation, but this does not invalidate the ion trap hierarchical keys for regio-isomer identification.(72)

Kuhnert *et al.* suggest that the reason these prototropic isomers can be resolved arises from them having significantly different cross sectional area and the phenolate ion being stabilised by hydrogen bonding after inversion of the carboxy equatorial chair, as illustrated below for 5-CQA. For 1-CQA and 3-CQA the hydrogen bonds favour the carboxy equatorial conformer and this may explain why only one prototropisomer was detected when 3-CQA was examined.



Carboxylate and phenolate ions of 5-caffeoylquinic acid, and hydrogen bond-stabilised phenolate ion from Kuhnert *et al.*(72) The hydrogen bond-stabilised inverted chair conformer yields a more intense fragment at m/z 191 relative to the parent ion at m/z 353.

Note, that the 1,4-dicaffeoyl-, 3,4-dicaffeoyl- and 4,5-dicaffeoyl-quinic acids differ markedly in terms of the secondary peaks in their MS² spectra, particularly the intensity of the dehydrated ion at *m/z* 335. This observation is useful when assigning other diacyl-quinic acid subgroups. It is essential when using the targeted mode of operation to assign tri-acyl-quinic acids, and chlorogenic acids incorporating aliphatic acids or having any substituent at C1.

Caffeoyl-feruloylquinic acids

Table 9. Base peak and selected secondary fragments for CFQA at MS², MS³ and MS⁴					
	MS¹	MS²	MS³	MS⁴	
3F,4CQA	529	353 + 367 (90%) 335 (75%) 349 (50%)	173 + 179 (90%)	134	
3C,4FQA	529	367 + 335 (20%) 349 (5%)	173	93	
3F,5CQA	529	367 + 335 (9%)	193	134	
3C,5FQA	529	353 + 367 (50%)	191 + 179 (60%)	127	
4F,5CQA	529	367	173 + 193 (60%)	93	
4C,5FQA	529	353 + 367 (25%)	173 + 179 (80%)	93	

Data are taken from Clifford *et al.* (14)

An examination of the MS² data for the six CFQA clearly identifies whether the first fragmentation step removes a caffeoyl or a feruloyl moiety. The MS³ fragmentation then defines the structure of the MS² ion and which hydroxycinnamic acid is present. Accordingly the first CFQA is easily assigned as 3F,4CQA because the MS² base peak is clearly [4-CQA]⁻, the feruloyl moiety was at either C3 or C5, and 3,4-diacyl-quinic acids consistently elute before 4,5-diacyl-quinic acids as discussed in the preceding section. Notice also that in the MS² spectra the dehydrated ions at *m/z* 335 and 349 are prominent for the 3,4-CFQA but not detectable for the 4,5-CFQA. The other five CFQA can be assigned by similar arguments and none of the assignments so obtained are in any way equivocal or conflicting. There are insufficient data for the 1-acyl CFQA but it should be possible to recognise them and possibly to assign them by their fragmentation supported by the use of a green robusta coffee bean extract as a surrogate standard.

Of potential importance is the use in these studies of either a triple-quadrupole-MS,(37) or a Brucker ion-trap-MS and operating conditions (36)(59) different from those used by the Kuhnert research group, and it is possible that this explains the different fragmentation of compounds that on the basis of their retention times appear to be acyl-quinic acid glycosides.

In principle, a full characterisation of dicaffeoylquinic acid glycosides ($m/z = 677.1718$) can be achieved by a targeted MS⁴ analysis at $m/z (677 + 515 + 353)$,(17, 75) However, there are two quite different possibilities for the identity of the fragment ion at $m/z 515$ — [dicaffeoylquinic acid]⁻ (accurate mass $m/z = 515.1189$), and [caffeoylquinic acid-glycoside]⁻ (accurate mass $m/z = 515.1402$). Only the [caffeoylquinic acid glycoside]⁻ can produce $m/z 341$, but either could give rise to $m/z 353$, and it is therefore not possible certainly to determine the origin of $m/z 353$ unless accurate mass is available.

Based on the known behaviour of caffeoylquinic acid glycosides discussed above, it is reasonable to assume that both ions will be produced by at least some isomers. If both are produced, both are trapped (regardless of whether or not accurate mass is available) and subsequently fragmented, and as a consequence it is not possible to determine which caffeic acid residue was lost first.

An added complication is the occurrence of caffeoylquinic acid diglycosides (biosides) ($m/z = 677.1929$), one of which has also been reported in *Lonicera henryi*. The compound in question yielded $m/z 515$ at MS² indicating the loss of one sugar moiety and $m/z 173$ at MS³, indicating that the caffeic acid moiety was at C4.(75) This compound also produced a significant MS² secondary ion at $m/z 497$, an ion seen as the MS² base peak in two compounds reported in *Chrysanthemum*,(17) and originally thought to be dicaffeoylquinic acid glycosides, but possibly wrongly assigned. A report of an apparent caffeoylquinic acid diglycoside eluting at 63 min seems much more likely to be a tricaffeoylquinic acid.(77)

Further studies on a wider range of these glycosides are required to gain a better insight into their behaviour. Reducing the collision energy for the first fragmentation might help by producing only one of the two fragmentations that yield the same nominal mass.

Mammalian sulphate conjugates of acyl-quinic acids

Caffeoylquinic acid sulphates and feruloylquinic acid sulphates have been observed as human metabolites of chlorogenic acids. The sulphate conjugates elute well in advance of the free acids.(78, 79) The sulphate moiety is readily lost at MS¹ and subsequent MS² fragmentation can define the regio-chemistry as previously described for the chlorogenic acids.

behaviour of scopoletin varies significantly with the experimental conditions. Scopoletin might elute before (93) or after (94) 5-CQA IUPAC, and between 5-CQA and 4-CQA IUPAC when 4-CQA elutes after 5-CQA.(95) Although it has not been possible to locate LC-ion trap-MS fragmentation data for scopoletin under the conditions used to discriminate between the caffeoylquinic acid regio-isomers it is clear that scopoletin can also yield fragment ions at m/z 353, 191, 179, 173 and 161 and that there is a real risk of confusion.(94) Discrimination might be more easily achieved in positive ion mode because it has been reported that an ion at m/z 193 is diagnostic for the scopolin fragment following deglycosylation.(93, 94)

3.1.5. Distinguishing Acyl-quinic Acids not containing (-)-Quinic Acid (L-Quinic Acid) IUPAC

Table 14. Chlorogenic Acids not containing (-)-quinic Acid (L-Quinic Acid) IUPAC					
	MS ¹	MS ²	MS ³	MS ⁴	Reference
3-caffeoyl- <i>muco</i> -quinic acid	353	191	127 + 173 (70%) 85 (50%) 109 (25%) 145 (20%)	109	(50)
3-feruloyl- <i>muco</i> -quinic acid	367	191	127 + 173 (30%) 85 (25%) 109 (15%) 145 (15%)	109	(50)
Caffeoylquinic acid epimer from maté	353	191 + 179 (50%) 173 (99%) 135 (10%) See note 1 below	155 + 127 (28%) 171 (46%)		(47)
Caffeoylquinic acid epimer from maté	353	179 + 191 (51%) 161 (10%) 135 (20%)	135		(47)
Caffeoylquinic acid epimer from <i>Rudbeckia hirta</i>	353	191 + 295 (13%)	173 + 127 (15%) 93 (39%)		(5)
Caffeoylquinic acid epimer from <i>Helianthus tuberosus</i> and <i>Carlina acaulis</i>	353	191 + 295 (15%)	173 + 127 (36%) 85 (46%)		(5)

Note 1: the MS² ions at *m/z* 191 and 173 are of almost identical intensity and which appears as base peak will vary from analysis to analysis. It is convention in such cases to assign the larger fragment as the base peak.

There is no doubt that these caffeoyl- and feruloyl-quinic acids are distinct from those known to contain (-)-quinic acid IUPAC, and distinct from the acyl-isocitric acids discussed in the preceding section. Only the first two in Table 13, prepared synthetically, but found as products of roasting in roasted coffee beans, have been fully characterised. There are more for which MS data are not available — see Part 2 of these notes.

